PANM 13

Drahoslava Janovská; Vladimír Janovský
Remark on computing the analytic SVD

In: Jan Chleboun and Karel Segeth and Tomáš Vejchodský (eds.): Programs and Algorithms of Numerical Mathematics, Proceedings of Seminar. Prague, May 28-31, 2006. Institute of Mathematics AS CR, Prague, 2006. pp. 111–121.

Persistent URL: http://dml.cz/dmlcz/702826

REMARK ON COMPUTING THE ANALYTIC SVD*

Dáša Janovská, Vladimír Janovský

Abstract

A new technique for computing Analytic SVD is proposed. The idea is to follow branches for just one selected singular value and the corresponding left/right singular vector.

1. Introduction

A singular value decomposition (SVD) of a real matrix $A \in \mathbb{R}^{m\times n}$, $m \geq n$, is a factorization $A = U\Sigma V^T$, where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal matrices and $\Sigma = \mathrm{diag}(s_1, \dots, s_n) \in \mathbb{R}^{m\times n}$. The values s_i, $i = 1, \dots, n$, are called singular values. They may be defined to be nonnegative and to be arranged in nonincreasing order.

Let A depend smoothly on a parameter $t \in \mathbb{R}$, $t \in [a, b]$. The aim is to construct a path of SVD's

$$A(t) = U(t)\Sigma(t)V(t)^T\,, \tag{1}$$

where $U(t)$, $\Sigma(t)$ and $V(t)$ depend smoothly on $t \in [a, b]$. If A is a real analytic matrix function on $[a, b]$, then there exists *Analytic Singular Value Decomposition* (ASVD), see [1]: There exists a factorization (1) that *interpolates* classical SVD defined at $t = a$, i.e.

- the factors $U(t)$, $V(t)$, and $\Sigma(t)$ are real analytic on $[a, b]$;
- for each $t \in [a, b]$, both $U(t) \in \mathbb{R}^{m\times m}$ and $V(t) \in \mathbb{R}^{n\times n}$ are orthogonal matrices and $\Sigma(t) = \mathrm{diag}(s_1(t), \dots, s_n(t)) \in \mathbb{R}^{m\times n}$ is a diagonal matrix;
- at $t = a$, the matrices $U(a)$, $\Sigma(a)$ and $V(a)$ are the factors of the classical SVD of the matrix $A(a)$.

Diagonal entries $s_i(t) \in \mathbb{R}$ of $\Sigma(t)$ are called *singular values.* Due to the requirement of smoothness, singular values may be negative and also their ordering may be arbitrary. Under certain assumptions, ASVD may be uniquely determined by the factors at $t = a$. For a theoretical background, see [9]. As far as the computation is concerned, an incremental technique is proposed in [1]: Given a point on the path,

*The research of both authors was partially supported by the Grant Agency of the Czech Republic (grant No. 201/06/0356). The first and the second author acknowledge financial support by the research projects MSM 6046137306 and MSM 0021620839, respectively, of The Ministry of Education, Youth and Sports, Czech Republic.

one computes a classical SVD for a neighboring parameter value. Next, one computes permutation matrices which link the classical SVD to the next point on the path. The procedure is approximative with a local error of order $O(h^2)$, where h is the step size.

An alternative technique for computing ASVD is presented in [12]: A non-autonomous vector field $H : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$ of a huge dimension $N = n + n^2 + m^2$ can be constructed in such a way that the solution of the initial value problem for the system $x' = H(t, x)$ is linked to the path of ASVD. Moreover, [12] contributes to the analysis of *non-generic points*, see [1], of the ASVD path. These points could be, in fact, interpreted as singularities of the vector field H. In [11], both approaches are compared.

A continuation algorithm for computing ASVD is presented in [7]. It follows a path of a few *selected* singular values and left/right singular vectors. It is aimed to treat large sparse matrices. The continuation algorithm is of a predictor-corrector type. The relevant predictor is based on Euler method hence on an ODE solver. In this respect, there is a link to [12]. Nevertheless, the Newton-type corrector guarantees the solution with a prescribed precision.

The continuation may get stuck at the points, where a nonsimple singular value $s_i(t)$ turns up for a particular parameter t and index i. In [1, 12], such points are called non-generic points of the path. They are related to the branching of singular value paths. The code in [7] incorporates extrapolation strategies in order to "jump over" such a point.

In the present contribution, we will review the continuation proposed in [7], see Section 2. We suggest and investigate the idea to continue just **one** singular value and the corresponding left/right singular vector. Finally, we report on numerical experiments.

2. Preliminaries

Let us recall the notion of a singular value of a matrix $A \in \mathbb{R}^{m\times n}$, $m \geq n$:

Definition 2.1 *We say that $s \in \mathbb{R}$ is a singular value of the matrix A if there exist $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$ such that*

$$Av - su = 0\,, \quad A^T u - sv = 0\,, \quad \|u\| = \|v\| = 1\,. \tag{2}$$

The vectors v and u are called the right and the left singular vectors of the matrix A.

Note that s is defined up to its sign: if the triplet (s, u, v) satisfies (2) then at least three more triplets

$$(s, -u, -v)\,, \quad (-s, -u, v)\,, \quad (-s, u, -v)\,,$$

can be interpreted as singular values, left and right singular vectors of A.

Definition 2.2 *For a given $s \in \mathbb{R}$, let us set*

$$\mathcal{M}(s) \equiv \begin{pmatrix} -sI_m & A \\ A^T & -sI_n \end{pmatrix},$$

where $I_m \in \mathbb{R}^{m\times m}$ and $I_n \in \mathbb{R}^{n\times n}$ are identities.

Definition 2.3 *We say that $s \in \mathbb{R}$ is a simple singular value of a matrix A if there exist $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$ such that*

$$(s, u, v), \quad (s, -u, -v), \quad (-s, -u, v), \quad (-s, u, -v)$$

are the only solutions to (2)*. A singular value s which is not a simple singular value is called nonsimple singular value.*

Remark 2.1 *Let $s \neq 0$.*

1. *s is a simple singular value of A if and only if* $\dim \operatorname{Ker} \mathcal{M}(s) = 1$.
2. *s is a simple singular value of A if and only if s^2 is a simple eigenvalue of $A^T A$. In particular, $v \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$,*

$$A^T A v = s^2 v, \quad \|v\| = 1, \quad u = \frac{1}{s} A v,$$

are the relevant right and left singular vectors of A.

Remark 2.2 *$s = 0$ is a simple singular value of A if and only if $m = n$ and* $\dim \operatorname{Ker} A = 1$.

Remark 2.3 *Let s_i, s_j, $s_i \neq s_j$, be simple singular values of A. Then $s_i \neq \pm s_j$.*

Let us recall the idea of [7]: The branches of *selected* singular values and corresponding left/right singular vectors $s_i(t)$, $U_i(t) \in \mathbb{R}^m$, $V_i(t) \in \mathbb{R}^n$ are considered i.e.,

$$A(t)V_i(t) = s_i(t)U_i(t), \quad A(t)^T U_i(t) = s_i(t)V_i(t), \tag{3}$$

$$U_i(t)^T U_i(t) = V_i(t)^T V_i(t) = 1 \tag{4}$$

for $t \in [a, b]$. The natural orthogonality conditions $U_i(t)^T U_j(t) = V_i(t)^T V_j(t) = 0$, $i \neq j$, $t \in [a, b]$, are added. Given p, $p \leq n$, the selected singular values $S(t) = (s_1(t), \ldots, s_p(t)) \in \mathbb{R}^p$, and the corresponding left/right singular vectors $U(t) = [U_1(t), \ldots, U_p(t)] \in \mathbb{R}^{m\times p}$, $V(t) = [V_1(t), \ldots, V_p(t)] \in \mathbb{R}^{n\times p}$ are followed as $t \in [a, b]$.

In the operator setting, let

$$F : \mathbb{R} \times \mathbb{R}^p \times \mathbb{R}^{m\times p} \times \mathbb{R}^{n\times p} \to \mathbb{R}^{m\times p} \times \mathbb{R}^{n\times p} \times \mathbb{R}^{p\times p} \times \mathbb{R}^{p\times p} \tag{5}$$

be defined as

$$F(t,X) \equiv \left(A(t)V - U\Sigma, A^T(t)U - V\Sigma, U^TU - I, V^TV - I\right) , \tag{6}$$

where $X \equiv (S,U,V) \in \mathbb{R}^p \times \mathbb{R}^{m\times p} \times \mathbb{R}^{n\times p}$, $\Sigma = \mathrm{diag}(S)$ and $I \in \mathbb{R}^{p\times p}$ is the identity. Under certain assumptions, the set of *overdetermined nonlinear equations*

$$F(t,X) = 0 \tag{7}$$

implicitly defines a curve in $\mathbb{R}\times\mathbb{R}^N$, where $\mathbb{R}^N$, $N = p(1+n+m)$, and $\mathbb{R}^p \times \mathbb{R}^{m\times p} \times \mathbb{R}^{n\times p}$ are isomorphic. The image of F, namely $\mathbb{R}^{m\times p}\times\mathbb{R}^{n\times p}\times\mathbb{R}^{p\times p}\times\mathbb{R}^{p\times p}$, and $\mathbb{R}^M$, $M = p(m+n+2p)$, are isomorphic.

The curve (7) can be parameterized by t, i.e. $t \mapsto X(t) = (S(t),U(t),V(t))$ so that $F(t,X(t)) = 0$ as $t \in [a,b]$. Given a solution $X(t)$ at $t = a$, the curve is initialized. For this purpose, we may select p singular values and left/right singular vectors computed via the classical SVD of the matrix $A(a)$, see e.g. [4].

In [7], the *tangent continuation*, see [2], Algorithm 4.25, p. 107, is applied. It is a predictor-corrector algorithm with an adaptive stepsize control. Let us note that the sparsity of $A(t)$ as $t \in [a,b]$ can be exploited.

3. Continuation of a single singular value

In this section, we will consider the idea of pathfollowing of **one** singular value and the corresponding left/right singular vector. We will expect the path to be locally a branch $s_i(t)$, $U_i(t) \in \mathbb{R}^m$, $V_i(t) \in \mathbb{R}^n$ satisfying conditions (3)&(4) for $t \in [a,b]$.

We consider the i-th branch, $1 \le i \le m$, namely, the branch which is initialized by $s_i(a)$, $U_i(a) \in \mathbb{R}^m$, $V_i(a) \in \mathbb{R}^n$ computed by the classical SVD, see [4]. Note that the SVD algorithm orders all singular values in descending order $s_1(a) \ge \ldots \ge s_i(a) \ge \ldots \ge s_m(a) \ge 0$. We assume that $s_i(a)$ is **simple**. For the analysis of this assumption, see Remark 2.1 and Remark 2.2.

Remark 3.1 *Let $s \neq 0$.*

1. *If $\mathcal{M}(s)\begin{pmatrix} u \\ v \end{pmatrix} = 0$ then $u^Tu = v^Tv$.*
2. *If in addition $\mathcal{M}(s)\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = 0$ then $u^T\tilde{u} = v^T\tilde{v}$.*
3. *s is a singular value of A if and only if $\dim\mathrm{Ker}\,\mathcal{M}(s) \ge 1$.*

For $\mathcal{M}(s)$, see Definition 2.2.

Note that if $s_i(t) \neq 0$ then due to Remark 3.1 one of the scaling conditions (4) is **redundant**. It motivates the following

Definition 3.1 *Consider a mapping*

$$f : \mathbb{R} \times \mathbb{R}^{1+m+n} \to \mathbb{R}^{1+m+n} \,,$$

$$t \in \mathbb{R} \,, \quad x = (s, u, v) \in \mathbb{R}^1 \times \mathbb{R}^m \times \mathbb{R}^n \longmapsto f(t, x) \in \mathbb{R}^{1+m+n} \,,$$

where

$$f(t, x) \equiv \begin{pmatrix} -su + A(t)v \\ A^T(t)u - sv \\ v^T v - 1 \end{pmatrix} . \tag{8}$$

As an alternative to (8) we will also use

$$f(t, x) \equiv \begin{pmatrix} -su + A(t)v \\ A^T(t)u - sv \\ u^T u + v^T v - 2 \end{pmatrix} \tag{9}$$

with an equivalent scaling.

The equation

$$f(t, x) = 0 \,, \quad x = (s, u, v) \,, \tag{10}$$

may locally define a branch $x(t) = (s(t), u(t), v(t)) \in \mathbb{R}^{1+m+n}$ of singular values $s(t)$ and left/right singular vectors $u(t)$ and $v(t)$. The branch is initialized at t^0 that plays the role of $t(a)$. It is assumed that there exists $x^0 \in \mathbb{R}^{1+m+n}$ such that $f(t^0, x^0) = 0$. The initial condition $x^0 = (s^0, u^0, v^0) \in \mathbb{R}^{1+m+n}$ plays the role of already computed SVD-factors $s_i(a) \in \mathbb{R}^1$, $U_i(a) \in \mathbb{R}^m$ and $V_i(a) \in \mathbb{R}^n$.

We solve (10) on an open interval $\mathcal{J}$ of parameters t such that $t^0 \in \mathcal{J}$.

Theorem 3.1 *Let $(t^0, x^0) \in \mathcal{J} \times \mathbb{R}^{1+m+n}$, $x^0 = (s^0, u^0, v^0)$ be a root of $f(t^0, x^0) = 0$. Assume that $s^0 \neq 0$ is a simple singular value of $A(t^0)$.*

Then there exists an open subinterval $\mathcal{I} \subset \mathcal{J}$ containing t^0 and a unique function $t \in \mathcal{I} \longmapsto x(t) \in \mathbb{R}^{1+m+n}$ such that $f(t, x(t)) = 0$ for all $t \in \mathcal{I}$ and that $x(t^0) = x^0$. Moreover, if $A \in C^k(\mathcal{I}, \mathbb{R}^{m\times n})$, $k \geq 1$, then $x \in C^k(\mathcal{I}, \mathbb{R}^{1+m+n})$. If $A \in C^\omega(\mathcal{I}, \mathbb{R}^{m\times n})$ then $x \in C^\omega(\mathcal{I}, \mathbb{R}^{1+m+n})$.

Proof Note that the assumptions yield that the partial differential $f_x(t^0, x^0) \in \mathbb{R}^{1+m+n} \times \mathbb{R}^{1+m+n}$ at (t^0, x^0) is a regular matrix.

Assuming $A \in C^k(\mathcal{I}, \mathbb{R}^{m\times n})$, $k \geq 1$, the statement is a consequence of Implicit Function Theorem, see e.g. [6]. In case that $A \in C^\omega(\mathcal{I}, \mathbb{R}^{m\times n})$, i.e. A is real analytic, again Implicit Function Theorem holds, see [10]. $\diamondsuit$

In case that $s^0 = 0$ is a simple singular value of $A(t^0)$, see Remark 2.2, the analysis is much more complicated. In the present paper we prefer to announce the result as a conjecture:

Conjecture 3.1 *Let* $(t^0, x^0) \in \mathcal{J} \times \mathbb{R}^{1+m+n}$, $x^0 = (s^0, u^0, v^0)$ *be a root of* $f(t^0, x^0) = 0$. *Assume that* $s^0 = 0$ *is a simple singular value of* $A(t^0)$ *i.e.* $m = n$ *and* $\dim \mathrm{Ker} A(t^0) = 1$. *Let* $(u^0)^T A'(t^0) v^0 \neq 0$.

Then there exists an open subinterval $\mathcal{I} \subset \mathcal{J}$ *containing* t^0 *and a unique function* $t \in \mathcal{I} \longmapsto x(t) \in \mathbb{R}^{1+2n}$ *such that* $f(t, x(t)) = 0$ *for all* $t \in \mathcal{I}$ *and* $x(t^0) = x^0$. *Moreover, if* $A \in C^k(\mathcal{I}, \mathbb{R}^{n\times n})$, $k \geq 1$, *then* $x \in C^k(\mathcal{I}, \mathbb{R}^{1+2n})$. *If* $A \in C^\omega(\mathcal{I}, \mathbb{R}^{n\times n})$ *then* $x \in C^\omega(\mathcal{I}, \mathbb{R}^{1+2n})$.

Let us compare:

Remark 3.2 *Consider the defining equation (7) for* $p = 1$. *It represents an* **overdetermined** *system for* $(t, X) \in \mathbb{R} \times \mathbb{R}^{1+m+n}$. *In [7], the condition (7) is meant in the least-squares sense. The compatibility of the solution set to (7), see [2] p. 93 for the notion, has been checked a posteriori. On the other hand, the formulation via (10) suggests that the solution set* (t, x) *to (10) is under certain assumption an implicitly defined curve in* $(t, x) \in \mathbb{R} \times \mathbb{R}^{1+m+n}$.

The practical advantage of (10) is that we can use the ready-made packages for continuation of an implicitly defined curves. In particular, we implemented a Matlab toolbox MATCONT, [3].

In Conclusions to [7], we admitted that the continuation of a bunch of p selected singular values and the relevant left/right singular vectors may get stuck. Note that the same phenomena was reported as the alternative methods are concerned, see [1, 12, 11]. In Introduction we noted that the continuation problems are related to nonsimple singular values on the path (see Definition 2.3). In [1, 12], these points are called *non-generic*.

Pathfollowing of the solution set of (10) via MATCONT is very robust. It does not usually get stuck. On the other hand, one has to be careful when interpreting the results. In principle, the minimal stepsize `MinStepsize` should be sufficiently small.

In [8], the non-generic points of the path are considered. The claim is that a non-generic point does not persist a sufficiently small perturbation of $A(t)$. In other words, given an $A(t)$ on a finite interval $a \leq t \leq b$ then, "usually", the set of non-generic points on the path is empty.

4. Numerical experiments

We consider the same problem as in [7] namely, the homotopy

$$A(t) = t\,A2 + (1-t)\,A1\,, \quad t \in [0,1]\,, \tag{11}$$

where the matrices

$$A1 \equiv \texttt{well1033.mtx}\,, \quad A2 \equiv \texttt{illc1033.mtx}$$

are taken over from `http://math.nist.gov/MatrixMarket/`. Note that $A1, A2 \in \mathbb{R}^{1033\times 320}$ are sparse while $A1$ and $A2$ are well and ill-conditioned. The aim is to continue

- 10 smallest singular values, left/right singular vectors of $A(t)$,
- 10 largest singular values, left/right singular vectors of $A(t)$.

The continuation is initialized at $t = 0$: The initial decomposition of $A1$ was computed via `SVDS`, see MATLAB Function Reference.

The results of continuation are resumed on Figure 1 and Figure 2. The branches are depicted in turns by solid and dash curves. This should underline that the branches do not cross each other. The computation complies with Theorem 3.1.

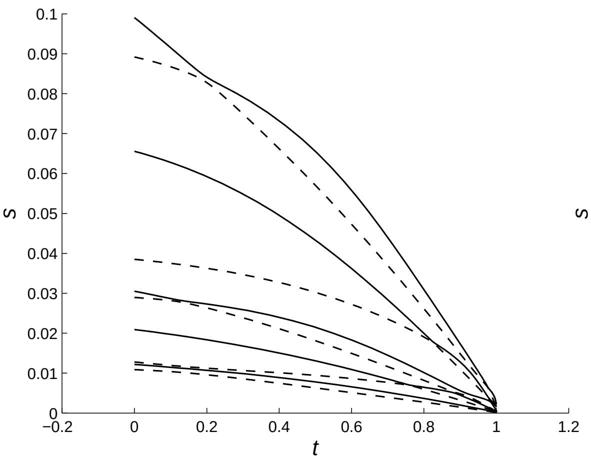

Fig. 1: *Ten smallest singular values s versus parameter t.*

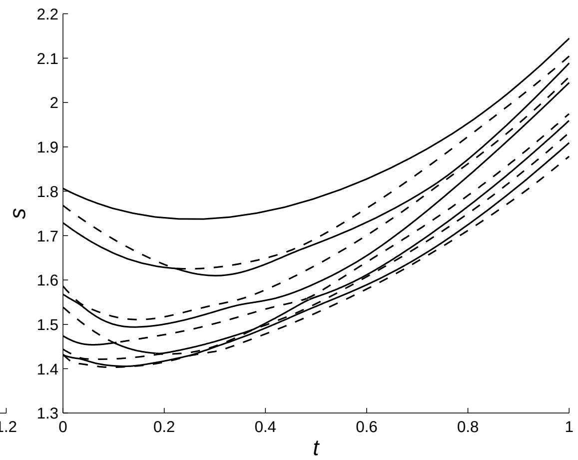

Fig. 2: *Ten largest singular values s versus parameter t.*

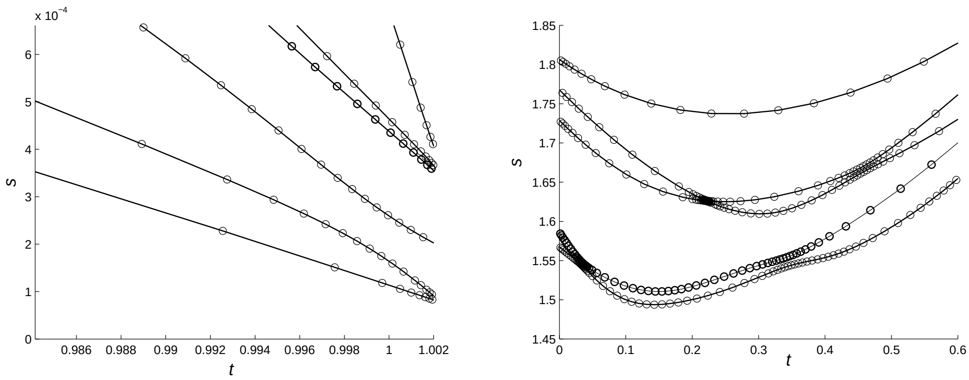

Fig. 3: *Zooms: Ten smallest s vs. t. Ten largest s vs. t.*

The zooms of the branches are shown on Figure 3. Each curve is computed as a sequence of isolated points marked by circles. The adaptive stepsize control refines the stepsize individually for each branch.

Note that the branches reported in [7] are not computed correctly. They cross each other occasionally: In the case of a stagnation, the continuation algorithm tries to jump over a prospective non-generic point on the path. A simple extrapolation strategy is used to continue. The branching scenario often suggests to follow a wrong branch. The message is that the branching is not generic.

In [7], the stepsize is always changed simultaneously for all p selected singular values. Treating each branch individually, see Figure 3, is much more efficient.

As the second example, we consider another homotopy

$$A(t) = t\, A3 + (1-t)\, A4\,, \quad t \in [-3, 10]\,, \tag{12}$$

where the matrices

$$A3 \equiv \texttt{cavity01.mtx}\,, \quad A4 \equiv \texttt{cavity02.mtx}$$

are taken over from `http://math.nist.gov/MatrixMarket/`. $A3, A4 \in \mathbb{R}^{317\times 317}$ are sparse square matrices.

The aim is to continue the smallest singular value and the relevant left/right singular vector over the interval $[-3, 10]$. The plot of the smallest singular value vs. t is shown on Figure 4. Note that $s(t)$ changes sign. It illustrates Conjecture 3.1: If the sign change occurs at t^0, the condition $(u^0)^T A'(t^0) v^0 \neq 0$ means that $s(t)$ crosses zero at t^0 "transversally", i.e. $s'(t^0) \neq 0$. It complies with the situation on Figure 4.

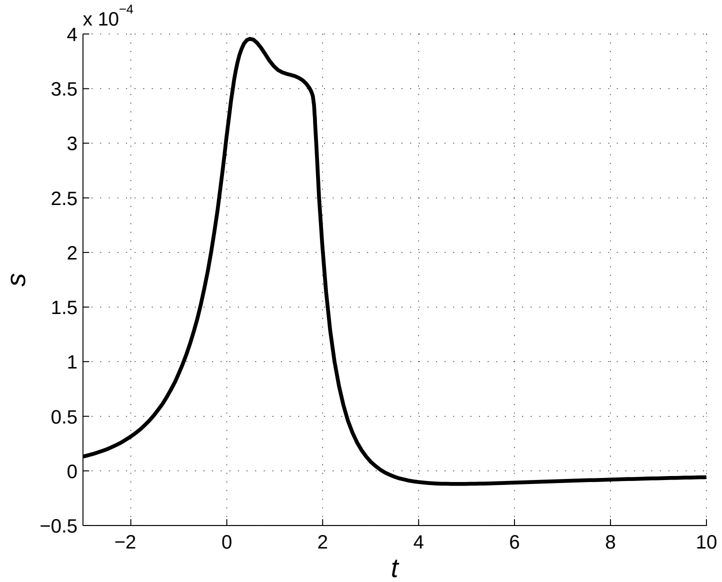

Fig. 4: *The smallest singular value s versus parameter t.*

5. Conclusions

In order to perform the Analytic SVD, we suggested to compute **separate** branches of singular values and the relevant left/right singular vectors. We can use any standard software for the pathfollowing of an implicitly defined curve. It seems, see [8], that the branches do not intersect generically. In other words, the branching scenario which concerns non-generic points, see [1, 12], does not persist sufficiently small perturbations of $A(t)$. So far, the claim is not rigorously proved. Nevertheless, the numerical experience supports the claim.

6. Appendix

We shall comment on Remark 3.1, Remark 2.1 and Remark 2.2. In particular, Remark 3.1 is based on Lemma 6.1 and Lemma 6.2, Remark 2.1 follows from Lemma 6.3 and Remark 2.2 is due to Lemma 6.4. Let us prove the Lemmas.

Lemma 6.1 *Let $s \neq 0$, $\mathcal{M}(s)\begin{pmatrix} u \\ v \end{pmatrix} = 0$. Then $u^T u = v^T v$.*

Proof By definition, we assume

$$-su + Av = 0\,, \quad A^T u - sv = 0\,.$$

Multiplying the first equation by u^T from the left and the second equation by v^T from the left, we get

$$u^T u = -\frac{1}{s} u^T A v\,, \quad v^T v = -\frac{1}{s} v^T A^T u\,.$$

Note that $v^T A^T u = (Av)^T u = u^T A v$. Therefore, $u^T u - v^T v = -\frac{1}{s}(u^T A v - u^T A v) = 0$.
$\diamondsuit$

Lemma 6.2 *Let $s \neq 0$, $\mathcal{M}(s)\begin{pmatrix} u \\ v \end{pmatrix} = 0$, $\mathcal{M}(s)\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = 0$. Then $u^T \tilde{u} = v^T \tilde{v}$.*

Proof We assume

$$-su + Av = 0\,, \quad A^T u - sv = 0\,,$$
$$-s\tilde{u} + A\tilde{v} = 0\,, \quad A^T \tilde{u} - s\tilde{v} = 0\,.$$

Therefore,

$$\tilde{u}^T(-su + Av) = 0\,, \quad \tilde{v}^T(A^T u - sv) = 0\,,$$
$$u^T(-s\tilde{u} + A\tilde{v}) = 0\,, \quad v^T(A^T \tilde{u} - s\tilde{v}) = 0\,.$$

Since $s \neq 0$,

$$\tilde{u}^T u = -\frac{1}{s}\tilde{u}^T A v\,, \quad \tilde{v}^T u = -\frac{1}{s}\tilde{v}^T A^T u = -\frac{1}{s}(A\tilde{v})^T u$$

and

$$u^T\tilde{u} = -\frac{1}{s}u^T A\tilde{v}\,, \quad v^T\tilde{v} = -\frac{1}{s}v^T A^T\tilde{u} = -\frac{1}{s}(Av)^T\tilde{u}\,.$$

We conclude that

$$\tilde{u}^T u - v^T\tilde{v} = -\frac{1}{s}\tilde{u}^T Av + \frac{1}{s}(Av)^T\tilde{u}\,.$$

Since $v^T\tilde{v} = \tilde{v}^T v$ and $(Av)^T\tilde{u} = \tilde{u}^T Av$,

$$\tilde{u}^T u - \tilde{v}^T v = 0\,.$$

$\diamondsuit$

Lemma 6.3 *A triplet $s \neq 0$, $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$ satisfies (2) if and only if*

$$A^T Av = s^2 v\,, \quad u = \frac{1}{s}Av\,, \quad \|v\| = 1\,, \quad s \neq 0\,. \tag{13}$$

Proof Let $s \neq 0$, u and v satisfy (2). From the first equation in (2), $0 = Av - su = 0$, we conclude that $0 = A^T(Av - su) = A^T Av - sA^T s = A^T Av - s^2 v$ since $A^T u = sv$. Moreover, $su = Av$, i.e. $u = \dfrac{1}{s}\,Av$.

Let $s \neq 0$, u and v satisfy (13). Then $A^T u - su = A^T(\frac{1}{s}Av) - sv = \frac{1}{s}A^T Av - sv = sv - sv = 0$ and $Av - su = Av - s(\frac{1}{s}Av) = Av - Av = 0$. Finally, $u^T u = u^T(\frac{1}{s}Av) = \frac{1}{s}u^T Av = \frac{1}{s}(A^T u)^T v = \frac{1}{s}\,sv^T v = 1$. $\diamondsuit$

Note that a nonzero simple singular value s can be identified with a nonzero simple eigenvalue s^2 of the matrix $A^T A$, see Lemma 6.3.

Lemma 6.4 *$s = 0$ is a simple singular value of A if and only if $m = n$ and* $\dim \mathrm{Ker}A = 1$.

Proof Let $m = n$, $\dim \mathrm{Ker}A = 1$. As a consequence, $\dim \mathrm{Ker}A^T = 1$. Then there exist $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^n$ such that

$$Av = 0\,, \quad A^T u = 0\,, \quad \|u\| = \|v\| = 1\,, \tag{14}$$

i.e. $(s = 0, u, v)$ satisfies (2). Clearly, $(s = 0, u, v)$ and $(s = 0, -u, -v)$ and $(s = 0, -u, v)$ and $(s = 0, u, -v)$ are the only possibilities to solve (2).

If $m > n$ then $\dim \mathrm{Ker}A^T \geq 2$ and hence (14) has infinitely many solutions. If $\dim \mathrm{Ker}A \geq 2$, one can also find infinitely many solutions to (14). $\diamondsuit$

References

[1] A. Bunse-Gerstner, R. Byers, V. Mehrmann, N.K. Nichols: *Numerical computation of an analytic singular value decomposition of a matrix valued function.* Numer. Math. **60**, 1991, 1–39.

[2] P. Deuflhart, A. Hohmann: *Numerical analysis in modern scientific computing. An introduction.* New York, Springer Verlag 2003.

[3] A. Dhooge, W. Govaerts, Yu.A. Kuzetsov: *MATCONT: A Matlab package for numerical bifurcation analysis of ODEs.* ACM Transactions on Mathematical Software **31**, 2003, 141–164.

[4] G.H. Golub, C.F. van Loan: *Matrix computations.* 3rd ed. Baltimore, The Johns Hopkins University Press 1996.

[5] W. Govaerts: *Numerical methods for bifurcations of dynamical equilibria.* Philadelphia, SIAM 2000.

[6] S.N. Chow, J.K. Hale: *Methods of bifurcation theory.* New York, Springer-Verlag 1982.

[7] V. Janovský, D. Janovská, K. Tanabe: *Computing the analytic singular value decomposition via a pathfollowing.* In: Proceedings of ENUMATH 2005, Springer Verlag, New York, 2006, 911–918.

[8] D. Janovská, V. Janovský: *On non-generic points of the analytic SVD.* In: T.E. Simos, G. Psihoyios, Ch. Tsitouras (Eds.), International Conference on Numerical Analysis and Applied Mathematics 2006., WILEY-VCH Verlag, Weinheim, 2006, 162–165.

[9] T. Kato: *Perturbation theory for linear operators, second ed.* New York, Springer Verlag 1976.

[10] S. Krantz, H. Parks: *A primer of real analytic functions.* New York, Birkhäuser 2002.

[11] V. Mehrmann, W. Rath: *Numerical methods for the computation of analytic singular value decompositions: electronic transactions on numerical analysis* **1**, 1993, 72–88.

[12] K. Wright: *Differential equations for the analytic singular value decomposion of a matrix.* Numer. Math. **63**, 1992, 283–295.